

OM protein - protein search, using sw model									
Run on: March 1, 2001, 15:52:46 ; Search time 170.72 Seconds (without alignments)									Copyright (c) 1993 - 2000 Compugen Ltd.
Post-processing: Maximum Match 0% Maximum Match 100% Listing first 45 summaries									GenCore version 4.5
Database : PIR:66:*									Accession: B24810
1: pir1:*									Score: 98.5
2: pir2:*									Score: 93.0
3: pir3:*									Score: 70.0
4: pir4:*									Score: 69.0
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.									
SUMMARIES									
Result No.	Score	Query	Length	DB	ID	Description	ALIGNMENTS		
1	98.5	28.7	639	2	B24810	beta-conglycinin alpha chain precursor - soybean	RESULT 1		
2	93.0	27.1	605	1	FWSYBA	beta-conglycinin alpha chain precursor - soybean	B24810		
3	93.0	27.1	605	2	S20007	beta-conglycinin alpha chain precursor - soybean	C:Species: Glycine max (soybean)		
4	75.0	21.9	411	2	T29475	beta-conglycinin alpha chain precursor - soybean	C:Accession: B24810; S16337		
5	70.0	20.4	221	2	T31863	beta-conglycinin alpha chain precursor - soybean	C:Accession: B24810; S16337		
6	69.0	20.1	1187	2	T18355	beta-conglycinin alpha chain precursor - soybean	C:Accession: B24810; S16337		
7	68.5	20.0	1513	2	T23681	beta-conglycinin alpha chain precursor - soybean	C:Accession: B24810; S16337		
8	67.5	19.7	438	1	T31889	beta-conglycinin alpha chain precursor - soybean	C:Accession: B24810; S16337		
9	67.5	19.7	445	2	T31898	beta-conglycinin alpha chain precursor - soybean	C:Accession: B24810; S16337		
10	66.5	19.4	509	2	S08059	alpha-globulin type III precursor	C:Accession: B24810; S16337		
11	66.5	19.4	771	2	D49423	alpha-globulin type III precursor	C:Accession: B24810; S16337		
12	66.0	19.2	242	2	T29699	alpha-globulin type III precursor	C:Accession: B24810; S16337		
13	66.0	19.2	388	2	T31887	alpha-globulin type III precursor	C:Accession: B24810; S16337		
14	66.0	19.2	388	2	T31888	alpha-globulin type III precursor	C:Accession: B24810; S16337		
15	66.0	19.2	605	2	S06398	alpha-globulin type III precursor	C:Accession: B24810; S16337		
16	64.0	18.7	1300	1	A36502	alpha-globulin type III precursor	C:Accession: B24810; S16337		
17	63.5	18.5	588	1	FWCNAB	alpha-globulin type III precursor	C:Accession: B24810; S16337		
18	63.5	18.5	151869	2		semaphorin III - m	C:Accession: B24810; S16337		
19	63.5	18.5	772	2	I48747	semaphorin III - m	C:Accession: B24810; S16337		
20	63.0	18.4	572	2	T29880	semaphorin III - m	C:Accession: B24810; S16337		
21	63.0	18.4	657	2	S32991	semaphorin III - m	C:Accession: B24810; S16337		
22	62.5	18.2	703	2	G6916	semaphorin III - m	C:Accession: B24810; S16337		
23	61.5	17.9	194	2	PN0667	semaphorin III - m	C:Accession: B24810; S16337		
24	61.0	17.8	151	2	S22145	semaphorin III - m	C:Accession: B24810; S16337		
25	61.0	17.8	181	2	T03740	semaphorin III - m	C:Accession: B24810; S16337		
26	61.0	17.8	545	1	A39827	semaphorin III - m	C:Accession: B24810; S16337		
27	61.0	17.8	545	1	B39827	semaphorin III - m	C:Accession: B24810; S16337		
28	61.0	17.8	728	2	T20561	semaphorin III - m	C:Accession: B24810; S16337		
29	60.5	17.6	263	2	T27641	semaphorin III - m	C:Accession: B24810; S16337		

A; Molecule type: mRNA  
 A; Residues: 1-605 <SEB>  
 A; Cross-references: EMBL:XI17698; NID:918535; PIDN:CAA35691.1; PID:918536  
 R; Shutov, A. D.; Kakhovskaya, T. A.; Bastrygina, A. S.; Bulmaga, V. P.; Horstmann, C.; Muentl, J. Bioclin. 241, 221-228, 1996  
 A; Title: Limited proteolysis of beta-conglycinin and glycinin, the 7S and 11S storage globulins  
 A; Reference number: S74123; MUID:97054613  
 A; Accession: S74124  
 A; Molecule type: protein  
 A; Residues: 189-196, 'H', 198, 'N', 200, 'X', 202-203; 397-408, 'X', 410, 'X', 412-417, 'X', 419-420.  
 A; Experimental source: seed  
 C; Superfamily: glycinin  
 C; Keywords: glycoprotein; seed; storage protein  
 F; 23-62/Domain: signal sequence #status predicted <PRO>  
 F; 63-605/Domain: propeptide #status predicted <MAT>  
 F; 261,517/Binding site: carbohydrate (Asn) (covalent) #status predicted  
  
 Query Match 27.1%; Score 93; DB 1; Length 605;  
 Best Local Similarity 31.4%; Pred. No. 0.011; 16; Indels 24; Gaps 3;  
 Matches 22; Conservative 8; Mismatches 16; Indels 24; Gaps 3;  
  
 Qy 2 ENPQAQRQLQSCQEPDLKQAKACESRTKLEYD-----PRCVYDGTGATNRP 50  
 Db 31 ENPKHNCQSCNSERDSYRNQACHARCNLLKVEKECCEGEIPRPR-----PRPQHP 83  
 Qy 51 -----PGER 54  
 Db 84 ERQQQQGK 93  
  
 RESULT 3  
 S20007  
 beta-conglycinin alpha chain precursor - soybean  
 C; Species: Glycine max (soybean)  
 C; Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 31-Oct-1997  
 C; Accession: S20007  
 R; Believre, J.-M.; Dickinson, C. D.; Dickinson, L.A.; Nielsen, N.C.  
 Plant Mol. Biol. 18, 259-274, 1992  
 A; Title: Synthesis and assembly of soybean beta-conglycinin in vitro.  
 A; Reference number: S20007; MUID:92119248  
 A; Accession: S20007  
 A; Status: preliminary  
 A; Molecule type: mRNA  
 A; Residues: 1-605 <LEL>  
 C; Superfamily: glycinin  
  
 Query Match 27.1%; Score 93; DB 2; Length 605;  
 Best Local Similarity 31.4%; Pred. No. 0.011; 16; Indels 24; Gaps 3;  
 Matches 22; Conservative 8; Mismatches 16; Indels 24; Gaps 3;  
  
 Qy 2 ENPQAQRQLQSCQEPDLKQAKACESRTKLEYD-----PRCVYDGTGATNRP 50  
 Db 31 ENPKHNCQSCNSERDSYRNQACHARCNLLKVEKECCEGEIPRPR-----PRPQHP 83  
 Qy 51 -----PGER 54  
 Db 84 ERQQQQGK 93  
  
 RESULT 4  
 T29475  
 hypothetical protein T01D1.6 - Caenorhabditis elegans  
 C; Species: Caenorhabditis elegans  
 C; Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 21-Jan-2000  
 C; Accession: T29475  
 R; Bradshaw, H.; Wohldmann, P.  
 Submitted to the EMBL Data Library, November 1996  
 A; Description: The sequence of C. elegans cosmid T01D1.  
 A; Reference number: 220523  
 A; Accession: T29475  
  
 RESULT 5  
 T31863  
 hypothetical protein C54F6.6 - Caenorhabditis elegans  
 C; Species: Caenorhabditis elegans  
 C; Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999  
 C; Accession: T31863  
 R; Geisel, C.; Bradshaw, H.  
 submitted to the EMBL Data Library, July 1997  
 A; Description: The sequence of C. elegans cosmid C54F6.  
 A; Reference number: Z21094  
 A; Accession: T31863  
 A; Status: preliminary; translated from GB/EMBL/DDBJ  
 A; Molecule type: DNA  
 A; Residues: 1-221 <GET>  
 A; Cross-references: EMBL:AF016447; PIDN:AAH65937.1; GSPDB:GN00023; CESP:C54F6.6  
 C; Genetics:  
 A; Experimental source: strain Bristol N2; clone C54F6  
 A; Gene: CESP:C54F6.6  
 A; Map position: 5  
 A; Introns: 53/1; 85/2; 123/3; 177/1.  
  
 Query Match 20.4%; Score 70; DB 2; Length 221;  
 Best Local Similarity 29.6%; Pred. No. 1.4; 16; Indels 16; Gaps 2;  
 Matches 16; Conservative 7; Mismatches 15; Indels 16; Gaps 2;  
  
 Qy 3 ENPQAQRQLQSCQEPDLKQAKACESRTKLEYD-----DPRQYD 41  
 Db 165 INNCAPKCFSSPDKKDENYLK-CETKCAKLTREMEEDDAKFSDFENCQACVYD 217  
  
 RESULT 6  
 T78355  
 hypothetical protein P3 - Mycoplasma hyorhinis  
 C; Species: Mycoplasma hyorhinis  
 C; Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 07-Dec-1999  
 C; Accession: T78355  
 R; Deng, G.; McIntosh, M.A.  
 J. Bacteriol. 176, 5929-5937, 1994  
 A; Title: An amplifiable DNA region from the Mycoplasma hyorhinis genome.  
 A; Reference number: Z18888; MUID:95014025  
 A; Accession: T18355  
 A; Status: preliminary; translated from GB/EMBL/DDBJ  
 A; Molecule type: DNA  
 A; Residues: 1-1187 <DEN>  
 A; Cross-references: EMBL:L11447; NID:9150156; PID:9664956; PIDN:AAA62228.1  
 C; Genetics:  
 A; Genetic code: SGC3

Matches 20; conservative 12; Mismatches 22; indels 16; Gaps 5; Result 9 T31898 hypothetical protein C03A7.14 - *Caenorhabditis elegans*

QY 1 TENPCQAQRQLQS--CQOPDDIKRQKAC---ESRCTKLEYDPRCV---YDT--GATN 46 C;Species: *Caenorhabditis elegans*

Db 403 TEER--ECTESTGCGCENPCECEBACDCSEBHCVCUDETQACLDCTNTOADTKVCGCTQ 460 C;Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 18-Feb-2000

C;Accession: T31898 R;Greco, T.; Bradshaw, H.; Elliott, G.

QY 47 ORHPPGERPR 56 C;Description: The sequence of *C. elegans* cosmid C03A7.

Db 461 EOHTCCEECK 470 C;Submitted to the EMBL Data Library, July 1997

submitted to the EMBL Data Library, November 1996 A;Description number: Z21096

A;Reference number: Z19781 A;Accession: T31898

A;Status: preliminary; translated from GB/EMBL/DDBJ A;Accession number: T31898

A;Molecule type: DNA A;Residues: 1-445 <GRE>

A;Cross-references: EMBL:AF016451; PIDN:AB66007.1; GSPDB:GN00023; CESP:C03A7.14

A;Experimental source: strain Bristol N2; clone C03A7

C;Genetics: A;Gene: CESP:C03A7.14

A;Map position: 5 A;Map position: 5

A;Introns: 1/8/3; 75/3 A;Introns: 1/8/3; 75/3

C;Superfamily: gliadin C;Superfamily: gliadin

Query Match 20.0%; Score 68.5; DB 2; Length 1513; Result 10 S08059 alpha-globulin type B precursor (tandem 1) - upland cotton (fragment)

Best Local Similarity 28.8%; Pred. No. 11; Matches 15; Conservative 9; Mismatches 11; Indels 17; Gaps 3; Matches 15; Conservative 7; Mismatches 16; Indels 7; Gaps 2; Result 10 S08059 alpha-globulin type B precursor (tandem 1) - upland cotton (fragment)

QY 4 PCQRCQCLSCQOE-----PDDIKRQKACESRCRCKLEYDPRCVYDTGATN 45 N;Alternate names: seed storage protein

Db 405 PCQPOCLOCSCLEHQIOPQVVTQLFQCIPO--COPAC-----EPQCIQETTT 449 C;Species: *Gossypium hirsutum* (upland cotton)

C;Accession: T31899 C;Accession: S08059 R;Chilan, C.A.; Borroto, K.; Kamalay, J.A.; Dure III, L.

R;Greco, T.; Bradshaw, H.; Elliott, G. Plant Mol. Biol. 9, 533-546, 1987

submitted to the EMBL Data Library, July 1997 A;Title: Developmental biochemistry of cottonseed embryogenesis and germination. XIX.

A;Description: The sequence of *C. elegans* cosmid C03A7. A;Residues: 1-448 <GRE> A;Residues: 1-448 <GRE>

A;Reference number: Z21096 A;Reference number: S08059

A;Accession: T31899 A;Accession: S08059

A;Status: preliminary; translated from GB/EMBL/DDBJ A;Status: not compared with conceptual translation

A;Molecule type: DNA A;Molecule type: DNA

A;Residues: 1-509 <CHL> A;Residues: 1-509 <CHL>

A;Cross-references: EMBL:AF016451; PIDN:AB65995.1; GSPDB:GN00023; CESP:C03A7.8 A;Cross-references: EMBL:AF016451; PIDN:AB65995.1; GSPDB:GN00023; CESP:C03A7.8

A;Experimental source: strain Bristol N2; clone C03A7 C;Genetics: A;Gene: CESP:C03A7.8

A;Map position: 5 A;Map position: 5

A;Introns: 75/3 A;Introns: 75/3

C;Superfamily: gliadin C;Superfamily: gliadin

Query Match 19.7%; Score 67.5; DB 2; Length 438; Result 11 D9423 semaphorin III precursor - human

Best Local Similarity 33.3%; Pred. No. 4.7; Matches 15; Conservative 7; Mismatches 16; Indels 7; Gaps 2; C;Species: *Homo sapiens* (man)

C;Accession: D49423 C;Accession: D49423 R;Kolodkin, A.L.; Matthes, D.J.; Goodman, C.S.

Cell 75, 1389-1399, 1993 A;Title: The Semaphorin genes encode a family of transmembrane and secreted growth co.

A;Accession: D49423 A;Accession: D49423 A;Status: preliminary; nucleic acid sequence not shown

A;Molecule type: mRNA A;Molecule type: mRNA

A;Residues: 1-771 <K0L> A;Residues: 1-771 <K0L>

Query Match 19.4%; Score 66; DB 2; Length 771;  
 Best Local Similarity 35.2%; Pred. No. 9.8; Mismatches 23; Indels 7; Gaps 2;  
 Matches 19; Conservative 5; Mismatches 23; Indels 7; Gaps 2;

Qy 6 AQRCQLQSQEQPDDLKQKACESRCTKLEYDPRCVYDGTATNQRHPPGERTRGRQ 59  
 Db 510 AQPLPLHRC-----DIYKACABCC--LARDPVCAMDSACSCSRYPTAKRIRRQ 556

RESULT 12  
 T39699 hypothetical protein F31A3.1 - *Caenorhabditis elegans*  
 C:Species: *Caenorhabditis elegans*  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 17-Mar-2000  
 C:Accession: T29699  
 R: Murray, J.; Le, T.  
 submitted to the EMBL Data Library, May 1996  
 A:Description: The sequence of *C. elegans* cosmid F31A3.  
 A:Reference: The sequence of *C. elegans* cosmid F31A3.  
 A:Accession: T29699  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA  
 A:Residues: 1-242 <MUR>  
 A:Cross-references: PIDN: AAB6656.1; GSPDB: GN00028; CESP: F31A3.1  
 A:Experimental source: strain Bristol N2; clone F31A3  
 C:Genetics:  
 A:Gene: CESP: F31A3.1  
 A:Map position: X  
 A:Introns: 18/3; 160/3  
 C:Superfamily: ultra-high-sulfur keratin

Query Match 19.2%; Score 66; DB 2; Length 242;  
 Best Local Similarity 43.3%; Pred. No. 4.1; Mismatches 13; Conservative 3; Mismatches 13; Conservative 3; Mismatches 12; Indels 2; Gaps 1;

Qy 2 EMPCAQRCLQSQEQPDDLKQKACESRCTK 31  
 Db 175 EPQCCQSCQOCVQOOSM0Q - CASACTK 202

RESULT 13  
 T31887 hypothetical protein C03A7.4 - *Caenorhabditis elegans*  
 C:Species: *Caenorhabditis elegans*  
 C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 21-Jan-2000  
 C:Accession: T31887  
 R: Graco, T.; Bradshaw, H.; Elliott, G.  
 Submitted to the EMBL Data Library, July 1997  
 A:Description: The sequence of *C. elegans* cosmid C03A7.  
 A:Reference number: Z21056  
 A:Accession: T31887  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA  
 A:Residues: 1-388 <GRE>  
 A:Cross-references: EMBL: AF016451; PIDN: AAB6596.1; GSPDB: GN00023; CESP: C03A7.7  
 A:Experimental source: strain Bristol N2; clone C03A7  
 C:Genetics:  
 A:Gene: CESP: C03A7.4  
 A:Map position: 5  
 A:Introns: 75/3  
 C:Superfamily: gliadin

Query Match 19.2%; Score 66; DB 2; Length 388;  
 Best Local Similarity 43.8%; Pred. No. 6.2; Mismatches 14; Conservative 4; Mismatches 8; Indels 6; Gaps 1;  
 Matches 23; Conservative 4; Mismatches 17; Indels 16; Gaps 5;

Qy 5 CAQRCQLQSQEQPDDLKQKACESRCT 30  
 Db 317 CAPOCEOSCOOCVQOOSPAOCQACQSCS 348

RESULT 14  
 T31888 hypothetical protein C03A7.7 - *Caenorhabditis elegans*  
 C:Species: *Caenorhabditis elegans*  
 C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 21-Jan-2000  
 C:Accession: T31888  
 R: Graco, T.; Bradshaw, H.; Elliott, G.  
 Submitted to the EMBL Data Library, July 1997  
 A:Description: The sequence of *C. elegans* cosmid C03A7.  
 A:Reference number: Z21096  
 A:Accession: T31888  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA  
 A:Residues: 1-388 <GRE>  
 A:Cross-references: EMBL: AF016451; PIDN: AAB6596.1; GSPDB: GN00023; CESP: C03A7.7  
 A:Experimental source: strain Bristol N2; clone C03A7  
 C:Genetics:  
 A:Gene: CESP: C03A7.7  
 A:Map position: 5  
 A:Introns: 75/3  
 C:Superfamily: gliadin

Query Match 19.2%; Score 66; DB 2; Length 388;  
 Best Local Similarity 43.8%; Pred. No. 6.2; Mismatches 14; Conservative 4; Mismatches 8; Indels 6; Gaps 1;  
 Matches 23; Conservative 4; Mismatches 17; Indels 16; Gaps 5;

Qy 9 CLQSC-----QEPDDLKQ-KACESRCTKLEYDPRCVYDGTATNQRHPPGERTRGRQ 54  
 Db 102 CEQSCREQYERQQQQPD--KQFKECQQRQCMQEPRKQCVKCREQYQEDWKG 159

Search completed: March 1, 2001, 15:52:47  
 Job time: 571 sec